



NLP 감정분석 기반 마케팅 시장 분석

빅데이터미네이터



목차

01

주제 고도화

- Input

02

선행 기술 조사

03

차별화

- Output

01

주제 고도화

- Input

주제 고도화

Input 구체화

“Input : YouTube 채널 댓글”

→ 도메인 : 쇼핑

→ 세부 주제 : 패션/의류 채널



주제 고도화

Input 구체화

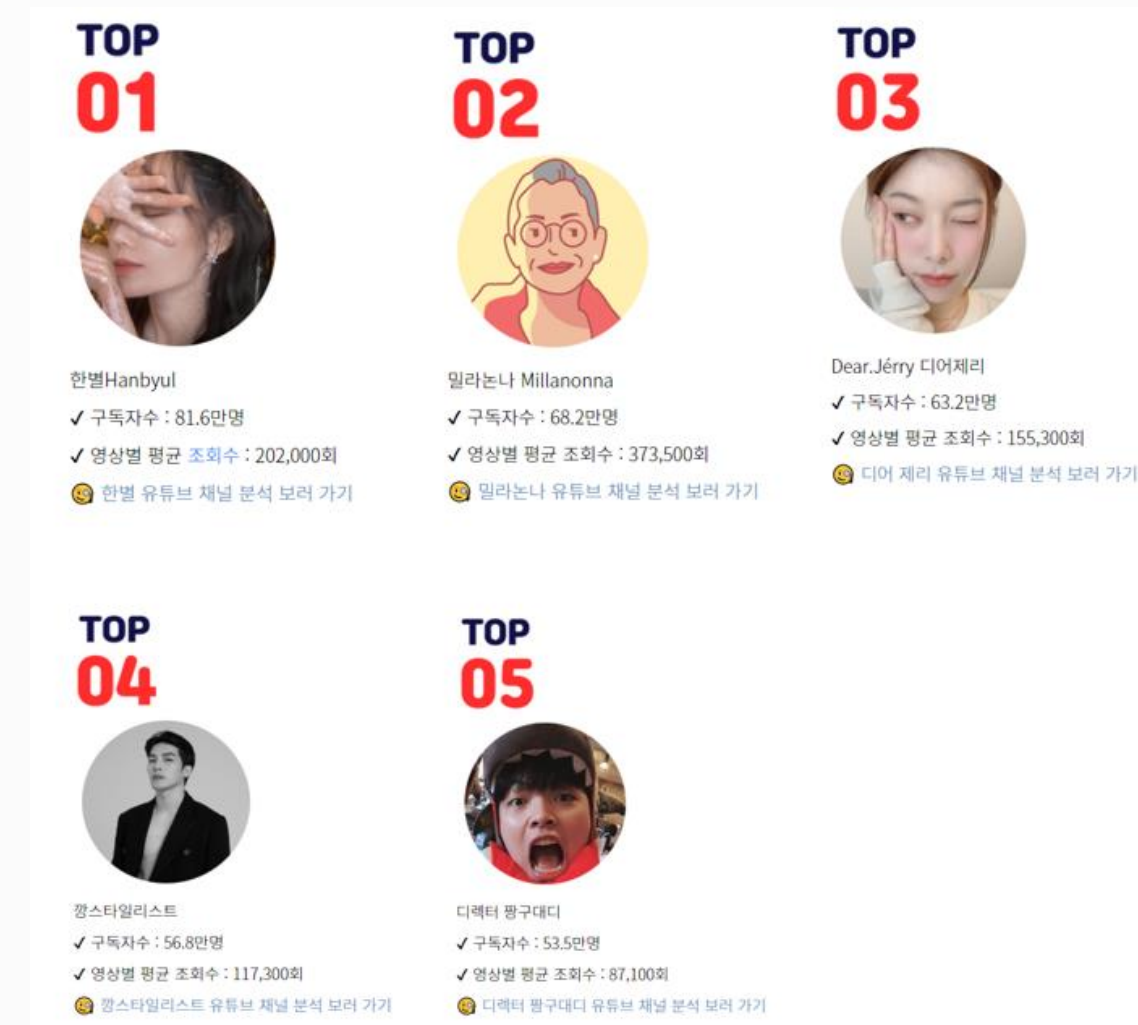
■ 참고할 만한 YouTube 채널

→ TOP 5 패션 유튜브 채널

(<https://www.i-boss.co.kr/ab-6141-52960>)

+ 그 외 참고할 만한 패션 유튜버

- 디자이너 or 스타일리스트 유튜버: '보라글레르', '옆집 언니 최실장'
- 여성 패션 유튜버: '헤인HEYNEE 's showwindow', 'AliceFunk 앨리스핑크'
- 남성 패션 유튜버: '핏더사이즈', '깡스타일리스트'




주제 고도화

Input 구체화

<채널 ‘밀라논나’ 댓글 분석>


→ 긍정 댓글

- 

KIM LUCY 1년 전

한섬 외국인 모델 좀 그만 썼음 좋겠어요 ㅠㅠㅠㅠ 시스템 옴르도 이 분이 입으신게 훨씬 낫고 옷발도 잘 사는거 같은데... 그리고 논나님 컬러매치 오늘도 배워갑니다 !!!

👍 2.2천 🗨️ 답글

▼ 답글 16개 보기
- 

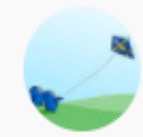
Cook Alexander 1년 전

어떤 분을 만나도 나이불문하고 항상 존중해주시는 태도가 너무 감동적이에요. 한국에서는 이렇게 대해주는 어르신 한 번도 못 만나봤는데.. 저도 나이들어서 논나님처럼 멋진 가치관과 태도를 습득한 어른이 될 수 있을까요

👍 1.2천 🗨️ 답글

▼ 답글 19개 보기

→ 부정 댓글


- 

무야 1년 전

192보고 그대로 돌아섭니다.ㅠㅠ

👍 🗨️ 답글

→ 추가 요청 사항 (기타)

- 

팡파르 fanfare 1년 전

여자편두 해주세용~ 몸에 걸러는 하나 내지 두 개만 !! 색깔매치 해서 옷 사는거 꿀팁이에요 저도 상하의 맞춰서 사면 그대로만 입거든요 😊 세탁법도 그렇구 언제나 도움되는 논나님 영상 최고예요 ❤️

👍 144 🗨️ 답글

분석 결과


➔ 인기 채널이고 큰 리스크가 없었기 때문에 대부분의 댓글이 긍정 반응인 편이다. 추가적으로 어떤 콘텐츠를 제작해 달라고 하는 요구 사항이 많이 올라오며, 부정의 정도가 심한 댓글은 많이 찾아보기 어렵다. 다만 이는 ‘밀라논나’ 채널의 경우 비교적 신규 채널이며, 주인공 '장명숙'님의 전문성이나 인성, 나이 등이 어느정도 참작된 경우이고, 타 대형 채널에서는 인기가 많은 만큼 '악성 댓글'에도 더 많이 시달리는 경우가 다수 있다고 한다.

주제 고도화

Input 구체화




<채널 '한별 Hanbyul' 댓글 분석>




이명인 4개월 전
내가 찼오래된 별별이로써 겨울마다 이 시리즈만 기다리지 진짜ㅠㅠㅠㅠㅠㅠㅠㅠ보기만해도 포근한 니트를 **책오...**💜💜

👍 135 🗨️ 🧡 답글




Scarlet Jung 4개월 전
진짜 한별님만큼 저랑 취향 비슷한 사람을 본적이 없어요 제가... 그래서 **사랑한다구요??**

👍 11 🗨️ 🧡 답글




Noelany Jen 4개월 전
초반 편집 센스 무엇ㅠㅠㅠㅠㅋㅋㅋ 이런 포인트 넘 **사랑한다구요오오** 가을 내놔ㅠㅠㅠㅠ 언니 ㄱ패션감각, 옷 선정 넘나 최고인 것...♥ 자주 보니까 넘 **조아요오오**♥♥♥

👍 51 🗨️ 🧡 답글



Rebecca 4개월 전
아 남편지갑이 어딴지 오늘도 지갑을 찾아 집안을 헤매는 레베카하이에나여 🐶
TMI 한별님 저희 세가족 아파트분양청약당첨되서 2년뒤에 이사가요😭😭 아가 태어나고 좋은 일만 생겨서 너무 **행복합니다**♥


→ 사랑, 행복 등은 분류하기 쉬운 긍정 단어이지만, 책오, 조아 등 변형된 한국어 분석하는 데 어려움이 있을 것이다. 이와 같은 비정형적인 텍스트를 어떻게 분석할 것인지 고민해 봐야 한다.



채원 4개월 전
언니 저 내년에 대학생 되면 언니가 추천해준 코디대로 입을꺼에오!!🥰🥰

👍 🗨️ 답글

→ 명확한 긍정적인 텍스트가 포함되어 있지 않지만, 문맥상 긍정적으로 인식되는 텍스트는 어떻게 분류할 것인지 생각해 봐야 한다.



을이 4개월 전
왜 저기 연보라색 니트는 안보여주시나요 ㅠㅠ 색갈 완전 제취향인데 ㅜㅜㅜ

👍 🗨️ 답글

→ 긍정, 부정도 아닌 애매한 텍스트는 어떻게 분석할 것인가? 종립으로 분류하는 것이 올바른가? "제 취향"은 긍정적인 단어로 볼 수 있지만 전체적인 문맥 상 긍정 반응으로 보기 어렵다.



02

선행 기술 조사

주요 선행 기술 문헌

01

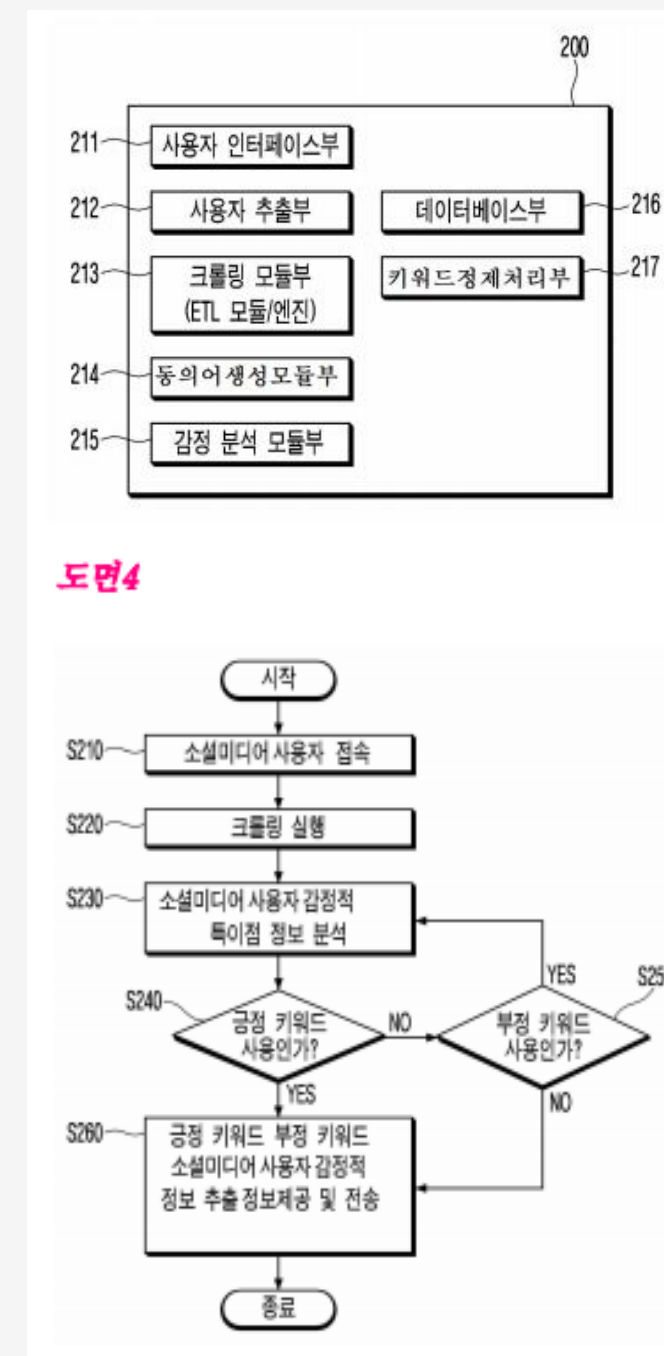
- 문헌번호 (공개 특허번호): 10-2020-0087357
- 기술 명칭: 빅데이터 분석을 통한 소셜 미디어 사용자 감정 분석 시스템 및 방법

- 기술 요약:

본 발명은, 빅 데이터 분석을 통한 소셜 미디어 사용자 감정 분석 시스템에 있어서, 소셜 미디어 상에서 공개된 정보를 이용하여 빅 데이터베이스를 생성하는 크롤링 모듈부, 상기 빅 데이터베이스 상에서 긍정 및 부정 반응에 대한 다양한 표현들의 동의어 사전을 생성하는 동의어 생성 모듈부, 및 상기 동의어 사전에 의하여 감정적 특이점 정보를 분석하는 감정 분석 모듈부를 포함할 수 있다.

- 기술 요소:

- 크롤링 기술 (크롤러 프로그램, 크롤링 엔진, ETL 모듈)
- 개인정보 필터링 시스템
- 데이터 정제 처리 시스템
- 관리 서버, 데이터베이스 구축



02

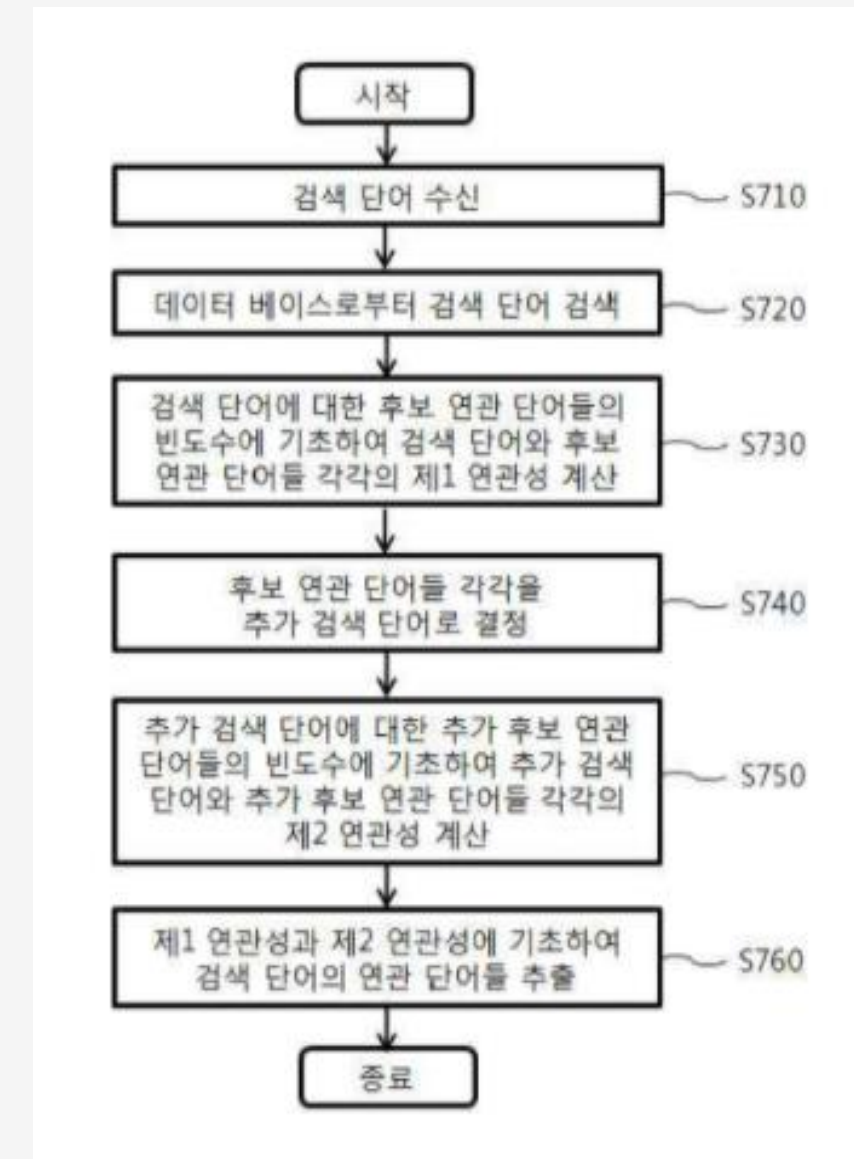
- 문헌번호: 10-2015-0053100
- 기술 명칭: 빅 데이터 처리 기술을 이용한 연관 단어 추출 방법 및 그 시스템

- 기술 요약:

본 발명은 연관 단어 추출 기술에 관한 것으로서, 빅데이터 처리 기술을 이용하여 검색 단어와 의미적 연관성이 높은 연관 단어들을 효율적으로 추출할 수 있는 연관 단어 추출 방법 및 그 시스템에 관한 것이다.

- 기술 요소:

- 데이터 수집 모듈
- 빈도수 추출 모듈
- 형태소 분석 모듈
- 데이터 저장 모듈
- 데이터 추출 모듈



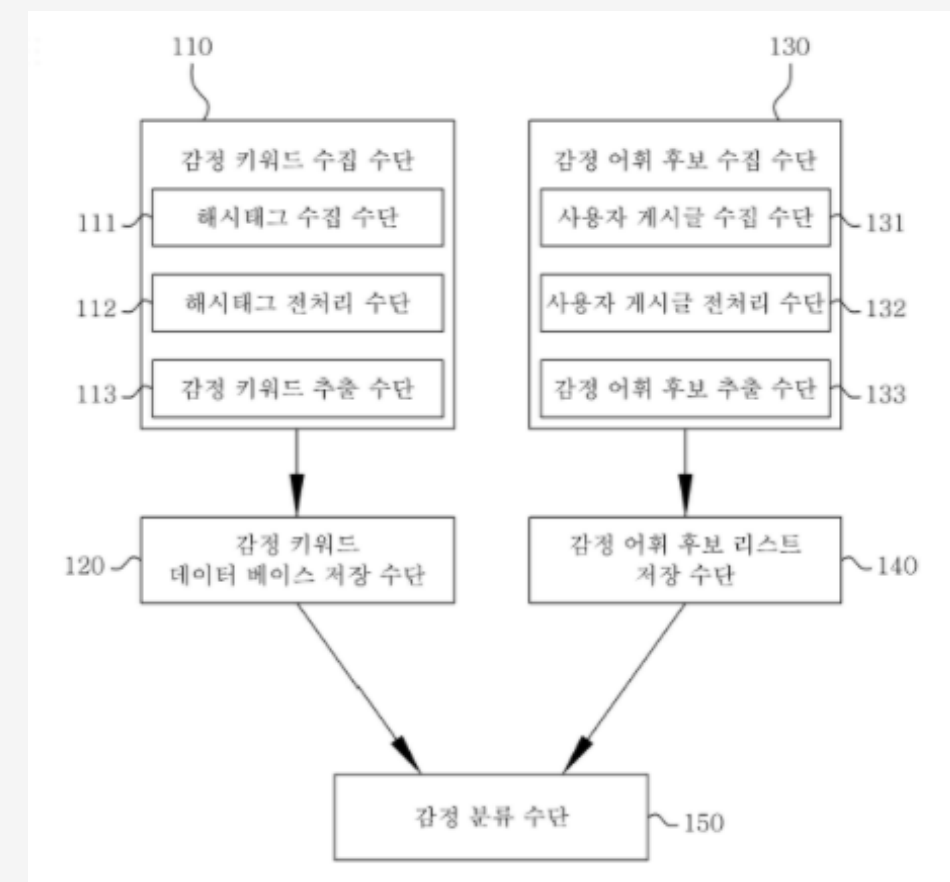
- 문헌번호: 10-2015-0151781
- 기술 명칭: 소셜 네트워크 서비스 상의 사용자 게시물 감정 분류 방법

• 기술 요약:

본 발명은 소셜 네트워크 서비스 상의 사용자 게시물 감정 분류 방법에 관한 것으로, 보다 구체적으로는 테이어 모델(Tayer Model)의 감정 카테고리로 구분되는 소셜 네트워크 서비스(Social Network Service) 상의 해시태그 (Hashtag)를 수집하여 감정 키워드 데이터 베이스를 구축하고, 사용자의 게시물과 감정 키워드 데이터 베이스의 각 감정 카테고리 간의 유사도를 비교하여 사용자 게시물의 감정 분포도 및 주요 감정을 객관적으로 분석하여 분류할 수 있는 소셜 네트워크 서비스 상의 사용자 게시물 감정 분류 방법에 관한 것이다.

• 기술 요소:

- 해시태그/사용자 게시물 수집 및 전처리 과정
- 감정 키워드(해시태그)/감정 어휘 후보(사용자 게시물) 추출 시스템
- 감정 키워드/감정 어휘 후보 데이터 베이스 구축
- 감정 분류 및 분석 시스템



04

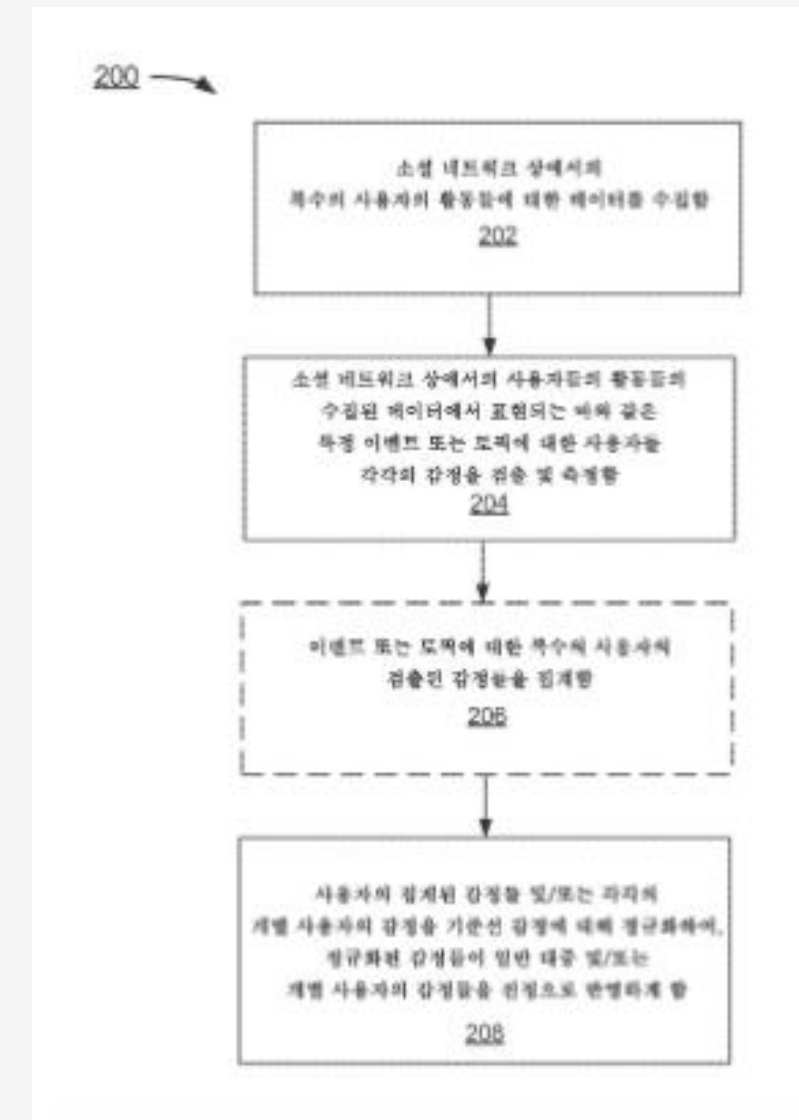
- 문헌번호 (공개 특허번호): 10-2014-0112008
- 기술 명칭: 소셜 네트워크 상에서의 감정 검출, 측정 및 정규화를 위한 시스템 및 방법

• 기술 요약:

소셜 네트워크 상에서 특정 이벤트 또는 토픽에 대해 사용자들의 그룹에 의해 표현되는 감정들을 검출, 측정, 집계 및 정규화하여, 정규화된 감정들이 그러한 특정 이벤트 또는 토픽에 대한 일반 대중의 감정들을 진정으로 반영하게 하는 능력을 제공하기 위한 시스템들 및 방법들을 고려하는 새로운 접근법이 제안된다. 게다가, 소셜 네트워크 상에서 표현되는 개별 사용자의 수집 및 측정된 감정들은 또한 사용자의 표현 시에 개별 사용자의 감정을 진정으로 반영하기 위해 반영하는 기준선 감정에 대해 정규화될 수 있다.

• 기술 요소:

- 데이터 크롤링 (텍스트 스코어링 스킴: sentiment text scoring scheme)
- 사용자들의 감정 정규화
- 감정 분석 엔진: 정규화된 측정에 소셜 감정 스코어(social sentiment score) 계산



+ 기타 참고 기술

①

- 문헌번호: 10-2020-0084190
- 특허권자: 주식회사 위팝
- 발명의 명칭: SNS 사용자의 소셜 미디어 특이점 정보 분석을 통한 맞춤형 마케팅 정보제공 시스템 및 방법

②

- 문헌번호: 10-1665649-0000
- 특허권자: 주식회사 엘지 씨엔에스
- 발명의 명칭: 소셜 미디어 데이터 분석 시스템 및 이를 이용한 소셜 미디어 데이터 분석 방법

③

- 문헌번호: 10-2017-0056160
- 특허권자: (주)아크릴
- 발명의 명칭: 딥러닝-기반 키워드에 연관된 단어를 도출하는 방법과 컴퓨터프로그램

③

- 문헌번호: 10-2019-0108522
- 특허권자: 숙명여자대학교 산학협력단
- 발명의 명칭: 댓글 기반의 동영상 광고효과 분석시스템 및 방법

* 제안 기술 요소

01

크롤링 - 추출
(전처리)

02

데이터베이스 -
감정어 사전 구축

03

감정분석 -
분류/분석/평가
(정제/정규화)

04

결과 도출 -
감정 분석 결과, 키워드 도출
(감정 키워드 / 요청 키워드)

선행기술 조사 종합 결론

- 종합적으로 선행 특허 조사를 진행한 결과 공개된 정보(댓글, 게시물 텍스트 등)를 이용하여 데이터베이스를 구축하고, 이를 활용해 미디어 사용자의 감정을 분석하는 빅데이터 분석을 통한 사용자 감정 분석 시스템 및 방법이라는 공통점이 있었다.
- 또한 빅데이터 처리 기술을 이용한 단어 추출 방법에 있어서, 상기 검색 단어와 상기 검색 단어에 대한 후보 연관 단어들 각각의 연관성을 계산하고, 연관 단어들을 추출한다는 공통점이 있었고, 수집한 데이터 각각의 개별 사용자 감정을 기준선 감정에 대해 정규화하여, 정규화된 감정들로 일반적인 반응을 분석한다는 점에서 데이터 정제 방식의 공통점이 있었다.
- 마지막으로 대다수 선행기술들은 소셜 네트워킹 플랫폼의 데이터를 분석/연구하였지만, 본 과제에서는 미디어 셰어링 플랫폼인 Youtube의 댓글만을 분석한다는 점에서 차이가 있었다.

+

03

차별화

- Output

주제 고도화

Output 구체화

→ Output 형태/수준:

프로그램의 결과 데이터를 제공하는 방식 (프로토타입)

(웹 서비스 확장성은 고민하는 중)

→ Output 차별화: 긍정 부정 외 구독자들의 요청 사항에 대한 키워드를 추출해서 제공

1. 다음 영상 콘텐츠에 대한 요청 키워드

: 긍정/부정에 관한 감정 분석 뿐 아니라, '주세요', '해주세요' 등이 포함된 요청 댓글에서 키워드를 추출하여 유튜버에게 제공해 준다.

유튜버는 제공받은 키워드를 바탕으로 차후 영상 콘텐츠에 관한 아이디어를 얻을 수 있고, 구독자들의 니즈를 세부적으로 파악할 수 있다.

2. 특정 제품에 대한 요청 키워드

: '정보좀', '정보', 등이 포함된 요청 댓글에서 키워드를 추출하여 유튜버에게 제공해 주면, 유튜버는 현재 유행하거나 수요가 많은

제품에 대한 인사이트를 얻을 수 있다.

mouji 2주 전

외국 옷 사이즈 도 해주세요

이번에cos 할인했는데 사이즈땀에 실패했어요 ㅠㅠ

하늘연달 2주 전

체형별 봉땀 2021버전 한번더 돌아와주세요♡♡

Eva Tomlinson 1주 전

실장님~ 목걸이 정보좀 알려주세요^^넘예빠요!



감사합니다 :))

빅데이터미네이터